UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
09/818,607	03/28/2001	Yasuo Okutani	35.G2761	1901

5514 7590 11/05/2004

FITZPATRICK CELLA HARPER & SCINTO
30 ROCKEFELLER PLAZA
NEW YORK, NY 10112

EXAMINER	
WOZNIAK, JAMES S	

ART UNIT	PAPER NUMBER
2655	

DATE MAILED: 11/05/2004

Please find below and/or attached an Office communication concerning this application or proceeding.

Office Action Summary	Application No.	Applicant(s)	
	09/818,607	OKUTANI ET AL.	
	Examiner	Art Unit	
	James S. Wozniak	2655	

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If the period for reply specified above is less than thirty (30) days, a reply within the statutory minimum of thirty (30) days will be considered timely.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

Status

1)☒ Responsive to communication(s) filed on *9/16/2004*.

2a)☐ This action is **FINAL**. 2b)☒ This action is non-final.

3)☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

Disposition of Claims

4)☒ Claim(s) *1,6,7,9-13,18,19 and 21-25* is/are pending in the application.

4a) Of the above claim(s) _____ is/are withdrawn from consideration.

5)☐ Claim(s) _____ is/are allowed.

6)☒ Claim(s) *1,6,7,9-13,18,19 and 21-25* is/are rejected.

7)☐ Claim(s) _____ is/are objected to.

8)☐ Claim(s) _____ are subject to restriction and/or election requirement.

Application Papers

9)☐ The specification is objected to by the Examiner.

10)☒ The drawing(s) filed on *28 March 2001* is/are: a)☒ accepted or b)☐ objected to by the Examiner.

Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).

Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).

11)☐ The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

Priority under 35 U.S.C. § 119

12)☒ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).

a)☒ All b)☐ Some * c)☐ None of:

1.☒ Certified copies of the priority documents have been received.

2.☐ Certified copies of the priority documents have been received in Application No. _____.

3.☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

* See the attached detailed Office action for a list of the certified copies not received.

Attachment(s)

1) ☒ Notice of References Cited (PTO-892)
2) ☐ Notice of Draftsperson's Patent Drawing Review (PTO-948)
3) ☐ Information Disclosure Statement(s) (PTO-1449 or PTO/SB/08) Paper No(s)/Mail Date _____.
4) ☐ Interview Summary (PTO-413) Paper No(s)/Mail Date. _____.
5) ☐ Notice of Informal Patent Application (PTO-152)
6) ☐ Other: _____.

DETAILED ACTION

Response to Amendment

1. In response to the office action from 8/17/2004, the applicant has submitted an amendment, filed 9/16/2004, amending Claims 1, 6, 7, 12, 13, 18, 19, and 24, while arguing to traverse the art rejection based on the limitation regarding a modification distortion as being between an individual unmodified synthesis unit and the same unit after modification *(Amendment, Page 9)*. Applicant's arguments have been fully considered, but are moot in view of the new ground of rejection with respect to Zinser *(U.S. Patent: 4,980,916)*.

Claim Objections

2. **Claim 25** is objected to under 37 CFR 1.75(c), as being of improper dependent form for failing to further limit the subject matter of a previous claim. Applicant is required to cancel the claim, or amend the claim to place the claim in proper dependent form, or rewrite the claim in independent form.

The infringement test for determining a proper dependent claim as per the MPEP 608.01 (n), Section III, states that a such a claim cannot conceivably be infringed by anything that would not also infringe the claim it references. In this case, a computer program product, such as a CD-ROM, would not infringe the method steps of Claim 13, 18, 19, or 21-24 since the program product *itself* never performs any of the active steps required by the claims. In other words

possession of such a program product would infringe Claim 20, but not Claims 13, 18, 19, or 21-24. Therefore, Claim 25 is an improper dependent claim.

Claim Rejections - 35 USC § 103

3. The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

> (a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negatived by the manner in which the invention was made.

4. **Claims 1, 12, 13, and 24** are rejected under 35 U.S.C. 103(a) as being unpatentable over Kagoshima et al *(U.S. Patent: 6,240,384)* in view of Zinser *(U.S. Patent: 4,980,916)*.

With respect to **Claims 1 and 13**, Kagoshima discloses:

Distortion obtaining means for obtaining a modification distortion between synthesis units before and after modification *(distortion calculator for determining a distortion between a synthesis speech segment and a training speech segment, Col. 13, Lines 58-60. Also, the training speech segment is modified with respect to pitch and duration to generate a synthesis speech segment, Col. 8, Lines 62-66)*.

Selection means for selecting synthesis units based on the modification distortion obtained by said distortion obtaining means *(selecting synthesis units that minimize distortion based on a distance comparison between synthesis and training units, Col. 2, Lines 58-62)*; and

Speech synthesis means for performing speech synthesis based on the synthesis units selected by said selection means *(speech synthesizer, Fig. 1, Element 15)*.

Kagoshima does not teach that the modification is obtained between an unmodified individual synthesis unit and the same individual unit after modification, however Zinser teaches a pitch error minimizer, which compares a pitch-altered synthesized speech sequence to an input or unmodified sequence to determine a distortion (error) *(Col. 3, Lines 5-37)*.

Kagoshima and Zinser are analogous art because they are from a similar field of endeavor in speech synthesis. Thus, it would have been obvious to a person of ordinary skill in the art, at the time of invention, to modify the teachings of Kagoshima with the means of determining a modification error as a difference between modified and unmodified versions of an individual synthesis units as taught by Zinser to improve synthesized speech quality for individual speech segments by selecting speech candidates for synthesis capable of minimizing a perceptual error *(Zinser, Col. 3, Lines 20-26)*.

With respect to **Claims 12 and 24**, Kagoshima further recites:

Input means and step for inputting text data *(input text, Col. 8, Line 10, that would inherently be inputted via a text input means)*;

Language analysis means and step for performing language analysis of the text data *(language processing of an input text, Col. 15, Lines 41-43)*; and

Prosody-parameter generation means and step for generating predetermined prosody parameters based on a result of analysis of said language analysis means and step *(obtaining prosody information from language processing, Col. 15, Lines 41-43)*.

Wherein said distortion obtaining means obtains the modification distortion based on the predetermined prosody parameters generated by said prosody parameter generation means *(distortion calculator for determining a distortion between a synthesis speech segment (training segment with added prosody information) and a training speech segment, Col. 13, Lines 58-60. Also, the training speech segment is modified with respect to pitch and duration to generate a synthesis speech segment, Col. 8, Lines 62-66, according to prosody information, Fig. 1, Element 111).*

5. **Claims 6 and 18** are rejected under 35 U.S.C. 103(a) as being unpatentable over Kagoshima et al in view of Zinser, and further in view of Huang et al *(U.S. Patent: 5,913,193).*

With respect to **Claims 6 and 18**, Kagoshima in view of Zinser teaches the speech synthesis apparatus and method that utilizes a modification distortion, calculated as the distance between an individual synthesis unit before and after modification, in selecting a best speech unit for synthesizing speech, as applied to Claims 1 and 13. Kagoshima in view of Zinser does not teach obtaining a distortion by adding modification and concatenation distortion, however Huang discloses:

A speech signal processing apparatus and method, wherein the distortion obtaining means uses a value obtained by adding the obtained modification distortion between the synthesis units before and after modification and a concatenation distortion *(spectral distortion between adjacent instances, Col. 3, Lines 1-6)* generated by concatenating a synthesis unit to another synthesis unit *(summing the distortions of an instance sequence, Col. 9, Lines 44-47).*

Kagoshima, Zinser, and Huang are analogous art because they are from a similar field of endeavor in speech synthesis. Thus, it would have been obvious to a person of ordinary skill in the art, at the time of invention, to modify the teachings of Kagoshima in view of Zinser with the method of summing distortions including a concatenation distortion as taught by Huang to further provide more natural synthesized speech by selecting a best synthesis unit dually based upon concatenation and modification distortion, thus minimizing distortion due to concatenation to create smooth transitions between speech units and modification to ensure natural sounding speech in the instance of a prosody change.

6. **Claims 7 and 19** are rejected under 35 U.S.C. 103(a) as being unpatentable over Kagoshima et al in view of Zinser, in further view of Huang et al, and in yet further view of Campbell et al *(U.S. Patent: 6,366,883)*.

With respect to **Claims 7 and 19**, Kagoshima in view of Zinser, and further in view of Huang teaches the speech synthesis system capable of selecting best speech instances based upon a concatenation and modification distortion sum, as applied to Claims 6 and 18. Kagoshima in view of Zinser, and further in view of Huang does not teach calculating a distortion as a weighted sum of modification and concatenation distortion, however Campbell discloses:

A speech signal processing apparatus and method, wherein the distortion obtaining means calculates a weighted sum of the modification distortion between the synthesis units before and after modification and the concatenation distortion generated by concatenating a synthesis unit to another synthesis unit *(selecting a speech unit based upon weighted coefficient vectors, Col. 2, Lines 37-38)*.

Kagoshima, Zinser, Huang, and Campbell are analogous art because they are from a similar field of endeavor in speech synthesis. Thus, it would have been obvious to a person of ordinary skill in the art, at the time of invention, to modify the teachings of Kagoshima in view of Zinser, and further in view of Huang with the method of selecting a speech unit based upon a weighted coefficient vector as taught by Campbell to provide a means of minimizing concatenation cost expressed through a weighting function and thus providing higher quality and audible synthesized speech.

7. **Claims 9 and 21** are rejected under 35 U.S.C. 103(a) as being unpatentable over Kagoshima et al in view of Zinser, and further in view of Campbell et al.

With respect to **Claims 9 and 21**, Kagoshima in view of Zinser teaches the speech synthesis apparatus and method that utilizes a modification distortion, calculated as the distance between an individual synthesis unit before and after modification, in selecting a best speech unit for synthesizing speech, as applied to Claims 1 and 13. Kagoshima in view of Zinser does not specifically suggest calculating modification distortion using a cepstrum distance, however Campbell discloses:

A speech signal processing apparatus and method, wherein said distortion obtaining means calculates the modification distortion using a cepstrum distance *(distortion calculation based upon prosodic feature parameters calculated from acoustic characteristics of speech units, namely, cepstral distance, Col. 12, Lines 1-36).*

Kagoshima, Zinser, and Campbell are analogous art because they are from a similar field of endeavor in speech synthesis. Thus, it would have been obvious to a person of ordinary skill

in the art, at the time of invention, to modify the teachings of Kagoshima in view of Zinser with the means of calculating distortion through cepstral distance as taught by Campbell to create a speech synthesis system in which modification distortion is calculated using cepstral distance, since cepstral distance is a specific example of the distance calculation taught by Kagoshima and a good way to describe a speech unit.

8. **Claims 10 and 22** are rejected under 35 U.S.C. 103(a) as being unpatentable over Kagoshima et al in view of Zinser, and further in view of Hon et al *(U.S. Patent: 6,490,563)*.

With respect to **Claims 10 and 22**, Kagoshima in view of Zinser teaches the speech synthesis apparatus and method that utilizes a modification distortion in selecting a best speech unit for synthesizing speech, as applied to Claims 1 and 13. Kagoshima in view of Zinser does not teach the use of a table to determine a distortion, however Hon discloses:

A speech signal processing apparatus and method, wherein the distortion obtaining means includes a table storing distortions, and determines the modification distortion by referring to the table *(use of a unit inventory that contains speech instances and a decision tree that denotes the best speech instances with regard to a joint distortion function consisting of a concatenation and prosody distortion, both of which may be stored in memory, Col. 6, Line 58- Col 7, Line 5)*.

Kagoshima, Zinser, and Hon are analogous art because they are from a similar field of endeavor in speech synthesis. Thus, it would have been obvious to a person of ordinary skill in the art, at the time of invention, to modify the teachings of Kagoshima in view of Zinser with the use of an inventory and decision tree denoting speech instances with respect to concatenation and prosodic distortion in selecting a best speech instance as taught by Hon to create a means of

saving distortion parameters for instances where similar text inputs exist- a stored distortion in an inventory and best instance saved in a decision tree could be looked up easily and be used for selecting the best speech instance, thus improving processing speed without degrading speech quality. It would also have been obvious to one of ordinary skill in the art, at the time of invention, to implement the inventory in a lookup table format, as is well known in the art, so that the speech unit with the least distortion could be selected.

With respect to **Claims 11 and 23**, Kagoshima in view of Zinser teaches the speech synthesis apparatus and method that utilizes a modification distortion, calculated as the distance between an individual synthesis unit before and after modification, in selecting a best speech unit for synthesizing speech, as applied to Claims 1 and 13. Kagoshima in view of Zinser do not teach the use of a table to determine a concatenation distortion, however Hon discloses:

A speech signal processing apparatus and method, wherein the distortion obtaining means includes a table storing distortions, and determines the modification distortion by referring to the table *(use of a unit inventory that contains speech instances and a decision tree that denotes the best speech instances with regard to a joint distortion function consisting of a concatenation and prosody distortion, both of which may be stored in memory, Col. 6, Line 58- Col 7, Line 5).*

Kagoshima, Zinser, and Hon are analogous art because they are from a similar field of endeavor in speech synthesis. Thus, it would have been obvious to a person of ordinary skill in the art, at the time of invention, to modify the teachings of Kagoshima in view of Zinser with the use of an inventory and decision tree denoting speech instances with respect to concatenation and prosodic distortion in selecting a best speech instance as taught by Hon to create a means of saving distortion parameters for instances where similar text inputs exist- a stored distortion in an

inventory and best instance saved in a decision tree could be looked up easily and be used for selecting the best speech instance, thus improving processing speed without degrading speech quality. It would also have been obvious to one of ordinary skill in the art, at the time of invention, to implement the inventory in a lookup table format, as is well known in the art, so that the speech unit with the least distortion could be selected.

9. **Claim 25** is rejected under 35 U.S.C. 103(a) as being unpatentable over Kagoshima et al in view of Zinser, Huang et al, Campbell et al, and in further view of Hon et al.

With respect to **Claim 25**, Kagoshima, Zinser, Huang, and Campbell in various combinations teach the method claims of 13, 18, 19, and 21-24. The aforementioned prior art does not specifically suggest method implementation using a storage medium, however, Hon discloses:

A storage medium, capable of being read by a computer, storing a program for executing a speech signal processing method *(computer readable storage medium containing computer instructions for implementing speech synthesis, Col. 4, Lines 36-39).*

Kagoshima, Zinser, Huang, Campbell, and Hon are analogous art because they are from a similar field of endeavor in speech synthesis. Thus, it would have been obvious to a person of ordinary skill in the art, at the time of invention, to modify the teachings of Kagoshima, Zinser, Huang, Campbell, and Hon with the use of a computer readable medium for implementing a speech synthesis method as taught by Hon to store a speech processing method on a computer readable medium to increase method compatibility and usability by providing a means for method use with multiple computer systems.

Conclusion

10. The prior art made of record and not relied upon is considered pertinent to applicant's disclosure:

- Amada et al *(U.S. Patent: 6,385,576)*- teaches a means for selecting a synthesis unit based upon a determined minimized error between input and modified synthesized speech.
- Miki et al *(U.S. Patent 5,396,576)*- teaches a method of selecting a segment of speech for synthesis based upon a distortion between input and synthesized speech.

11. Any inquiry concerning this communication or earlier communications from the examiner should be directed to James S. Wozniak whose telephone number is (703) 305-8669 and email is James.Wozniak@uspto.gov. The examiner can normally be reached on Mondays-Fridays, 8:30-4:30.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Doris To can be reached at (703) 305-4827. The fax/phone number for the Technology Center 2600 where this application is assigned is (703) 872-9306.

Any inquiry of a general nature or relating to the status of this application or proceeding should be directed to the technology center receptionist whose telephone number is (703) 306-0377.

SUSAN MCFADDEN
PRIMARY EXAMINER

James S. Wozniak
10/19/2004